JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, D20102, doi:10.1029/2009JD012038, 2009

Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product

César Coll,[1] Zhengming Wan,[2] and Joan M. Galve[1]

Received 11 March 2009; revised 21 July 2009; accepted 28 July 2009; published 17 October 2009.

[1] The V5 level 2 land surface temperature (LST) product of the Moderate Resolution Imaging Spectroradiometer (MODIS) was validated over homogeneous rice fields in Valencia, Spain, and the Hainich forest in Germany. For the Valencia site, ground LST measurements were compared with the MOD11_L2 product in the conventional temperature-based (T-based) method. We also applied the alternative radiance-based (R-based) method, with in situ LSTs calculated from brightness temperatures in band 31 through radiative transfer simulations using temperature and water vapor profiles and surface emissivity data. At the Valencia site, profiles were obtained from local radiosonde measurements and from National Centers for Environmental Prediction (NCEP) data. The R-based method was applied at the Hainich site using radiosonde profiles from a nearby sounding station and NCEP profiles. The T-based validation showed average bias (MODIS minus ground) of −0.3 K, standard deviation of 0.6 K and root mean square error (RMSE) of ±0.7 K. For the R-based method, the quality of the atmospheric profiles was assessed through the difference $\delta(T_{31}-T_{32})$ between the actual MODIS and the profile-based calculated brightness temperature difference in bands 31 and 32. For the cases where $-0.3\ K < \delta(T_{31}-T_{32}) < 0.5\ K$, the R-based method yielded LST errors with small biases and RMSE = ±0.6 K for the two sites. These results show the high accuracy and precision of the MODIS LST product for the two sites studied. The good performance of the R-based method opens the possibility for a more complete validation including heterogeneous surfaces where the T-based method is not feasible.

Citation: Coll, C., Z. Wan, and J. M. Galve (2009), Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product, *J. Geophys. Res.*, *114*, D20102, doi:10.1029/2009JD012038.

1. Introduction

[2] Land surface temperature (LST) is a parameter resulting from the physical interactions in the surface-atmosphere system, particularly the energy and water fluxes between the atmosphere and the ground. LST is sensitive to local atmospheric conditions, surface type, soil humidity and vegetation water stress, among others, thus being a key variable in meteorological, climatological and hydrological applications. LST products retrieved from the Moderate Resolution Imaging Spectroradiometers (MODIS) [*Salomonson et al.*, 1989] on the Terra (morning) and Aqua (afternoon) platforms have been made available in the current version 5 (V5). *Wan* [2008] presented details of the new refinements made in the V5 MODIS LST algorithms. The day/night LST algorithm [*Wan and Li*, 1997] was implemented through a full incorporation of the split window algorithm into the day/night algorithm so that the LSTs in the resulting LST/emissivity products MOD11B1 (Terra) and MYD11B1 (Aqua) at 6-km grids are closer to the LSTs retrieved by the split window algorithm. However, the same generalized split window algorithm [*Wan and Dozier*, 1996] has been used to generate the V4 and V5 level 2 LST products MOD11_L2 (Terra) and MYD11_L2 (Aqua) at 1-km resolution. M*D is used for both Terra (MOD) and Aqua (MYD) products. The difference between V4 and V5 M*D11_L2 products due to the small differences in the V4 and V5 input data (especially the calibrated radiance, geolocation and atmospheric profile products) is less than 0.2 K in most cases.

[3] The retrieval of LST from satellite data involves the correction of the satellite-observed radiances for atmospheric absorption and emission and nonunity of land surface emissivity. It is therefore necessary to assess the accuracy and precision of the retrievals to provide potential LST users with reliable information on the quality of the data. Long-term validation is required to identify possible deficiencies and subsequently introduce improvements in the algorithms. A conventional method for satellite LST validation is the direct comparison with ground measurements performed at a field site concurrently with the satellite overpass. This method is referred to as temperature-based (T-based). It is necessary that the LST observed by the ground instruments at several points over the test site is truly representative of the average LST over the instantaneous field of view of the satellite sensor, for which the site must be thermally homogeneous from the point scale to several kilometers.

[1] Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, Burjassot, Spain.
[2] ICESS, University of California, Santa Barbara, California, USA.

0148-0227/09/2009JD012038$09.00

Since most of the Earth's surface is heterogeneous at these spatial scales, high-quality ground validation data are limited to few surface types such as lakes, silt playas, grasslands and agricultural fields collected during dedicated campaigns [*Wan et al.*, 2002, 2004; *Coll et al.*, 2005]. An exception is the Lake Tahoe automated validation site [*Hook et al.*, 2007], where lake surface temperatures are continuously measured since 1999. Recently, automated LST validation sites have been established in Europe and Africa [*Trigo et al.*, 2008].

[4] An alternative for the validation of LST products is referred to as the radiance-based (R-based) method [*Wan and Li*, 2008]. It does not require in situ measurements of LST. Instead, atmospheric profiles of temperature and water vapor over the site at the time of the satellite overpass are necessary, together with measurements or estimations of the surface emissivity. Basically, the method consists of calculating the ground LST from the top of the atmosphere (TOA) brightness temperature using the emissivity and atmospheric profile data in a radiative transfer model. Since sites with small variations in surface emissivity are more frequent than homogeneous areas in LST, the R-based method can be potentially applied to a larger number of sites than the T-based method.

[5] This paper shows results of both T-based and R-based validations of the MODIS M*D11_L2 (V5) LST product. The T-based method was applied using ground data from an agricultural site in Valencia, Spain. The test site was first described by *Coll et al.* [2005], where the Terra MOD11_L2 (V4) product was validated using ground data from 2002 to 2004. In this paper, we show the T-based validation for a total of 23 ground LST concurrences including 16 new cases from the 2005–2007 campaigns.

[6] In addition, we present R-based validation results for the same test site. For eight cases, we used atmospheric profiles measured at the site concurrently with Terra overpasses. Since it is difficult in practice to have local radiosonde measurements, we also used atmospheric profiles obtained from National Centers for Environmental Prediction (NCEP) data [*Kalnay et al.*, 1996] for the 23 above cases plus an additional data set of 51 Terra cases. The R-based method was further applied over the Hainich forest site in Germany, belonging to the CarboEurope Integrated Project (CarboEurope-IP, http://www.carboeurope.org). For this site, we selected 34 Terra and 16 Aqua scenes from March to October 2004. Atmospheric profiles were obtained from a radio sounding station in Meiningen, Germany, about 60 km south of Hainich, and from NCEP data as well.

[7] The paper follows with the description of the T-based and R-based methods. Section 3 shows the experimental data used in this study. Then, we discuss the validation results of the T-based method (section 4.1) and the R-based method (Valencia, section 4.2; Hainich, section 4.3). Finally, the main conclusions of the study are given in section 5.

2. T-Based and R-Based Validation

[8] The T-based method compares ground measured LSTs with satellite derived LSTs. Thus, its main advantage is that it provides a direct evaluation of the radiometric quality of the satellite instrument together with the ability of the LST retrieval algorithm to correct for atmospheric and emissivity effects. However, T-based validation depends critically on the accuracies of the ground measurements and how well they represent the surface temperature over which they were performed. This is a problematic issue due to the difficulty in scaling up from the ground point measurements to the field of view of the satellite sensor [*Wan et al.*, 2002]. During the day, LSTs can vary by 10 K or more over a few meters depending on the nature of the surface and the local meteorological conditions, the variability being lower at night. For this reason, only thermally homogeneous sites at the various spatial scales of the validation process must be used.

[9] The thermal homogeneity of potential validation sites can be assessed using thermal infrared (TIR) satellite data at different spatial resolutions, e.g., MODIS data at 1 km, and Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) or Landsat data at ~100 m. At the point scale (~1 m), the thermal homogeneity of the area must be assessed by means of multiple (spatial and temporal) sampling of the ground LSTs. The most useful in situ LST data are obtained from TIR radiometric measurements since satellite LST data are representative of the skin temperature. Several ground radiometers (calibrated and intercalibrated at the field working conditions) must be distributed within the 1 km^2 grid to account for the natural variability of ground LSTs. Radiometric temperatures must be corrected for surface emissivity and the downwelling sky irradiance. Together with the in situ LST, an estimation of the overall error budget should be made, including the instrumental error, the natural variability of ground LSTs, and the emissivity correction error. Uncertainties better than ±1 K are required for useful LST validation studies. Such accuracies can only be achieved over a few surface types such as inland waters, silt playas and fully vegetated surfaces, usually in dedicated, short-term campaigns with limited LST and climatic regimes [*Wan et al.*, 2002, 2004; *Coll et al.*, 2005; *Hook et al.*, 2007].

[10] The R-based method [*Wan and Li*, 2008] provides an alternative for the global validation of MODIS LST products since it does not rely on the ground measured LST. The in situ temperature is estimated from the MODIS TOA radiance using spectral emissivity data for the selected surface, nearly concurrent atmospheric temperature and water vapor profiles, and an atmospheric radiative transfer code. The difference between the product and the calculated in situ LST is the accuracy of the MODIS LST product. MODIS band 31 (10.78–11.28 μm) is used in the in situ LST calculation because the effect of variations in atmospheric water vapor and temperature profiles is smallest in this spectral range, as well as the uncertainty in surface emissivity. Band 32 (11.77–12.27 μm) can also be used, although it is more sensitive to atmospheric variations. As for the T-based method, field emissivity measurements or estimations are required. Emissivity in bands 31 and 32 is high and exhibits small variations for most land cover types [*Snyder et al.*, 1998], so the R-based method could be applied, with care, globally.

[11] The R-based method depends on the accuracy of the radiative transfer model used. In this paper, we used the MODTRAN 4 code [*Berk et al.*, 1999] since it is publicly available and widely used by the remote sensing community. The accuracy of MODTRAN in the TIR region is discussed by *Merchant and Le Borgne* [2004] and *Dash and*

Ignatov [2008]. In case the model has biases, still the R-based method can be used for monitoring relative performances over areas where T-based validation is not feasible.

[12] The strongest limitation of the R-based method is the need for accurate atmospheric profiles, a problem that can be mitigated by selecting suitable validation areas close to meteorological sounding stations, or by using atmospheric data from NCEP global tropospheric reanalysis products. Both possibilities were investigated in this study. Additionally, we checked the accuracy of the atmospheric profiles using the difference $\delta(T_{31}-T_{32})$ between the actual brightness temperature difference in MODIS bands 31 and 32 and the $T_{31}-T_{32}$ value calculated from the atmospheric profiles and the surface emissivity data [*Wan and Li*, 2008].

[13] The $\delta(T_{31}-T_{32})$ test can be used as a quality check for the atmospheric profiles since $\delta(T_{31}-T_{32})$ should be close to zero when the atmospheric temperature and water vapor profiles used in simulations represent the real atmospheric conditions of the MODIS observations and the effect of the surface emissivity uncertainties in bands 31 and 32 is small. The test relies on the fact that the atmospheric effect is larger in band 32 than in band 31, owing to the water vapor continuum absorption. Then $T_{31}-T_{32}$ is usually positive and increases with the atmospheric water vapor. When the atmospheric profile used for the R-based LST calculation is over correcting the atmospheric effect, then $\delta(T_{31}-T_{32})<0$ since the calculated, profile based $T_{31}-T_{32}$ value is larger than the actual MODIS value. Similarly, when the atmospheric effect is underestimated, the profile based $T_{31}-T_{32}$ is smaller than the actual MODIS value and thus $\delta(T_{31}-T_{32})>0$. Therefore, accurate atmospheric profiles are associated with $\delta(T_{31}-T_{32})$ values in a narrow range around zero. Since band 32 is not used in the calculation of the R-based LST, it provides an independent means to check the suitability of the atmospheric water vapor and temperature profiles. The R-based method can be summarized as follows:

[14] 1. In situ LST ($T_{R\text{-based}}$) from band 31 TOA radiance is calculated using atmospheric profiles and surface emissivity data (inverse simulation). The LST error (δT) is the difference between the product LST and $T_{R\text{-based}}$.

[15] 2. Brightness temperatures in bands 31 and 32 are calculated using $T_{R\text{-based}}$ as ground LST, atmospheric profiles and surface emissivity data (direct simulation). The difference $\delta(T_{31}-T_{32})$ between the actual and the simulated $T_{31}-T_{32}$ value is obtained.

[16] A good knowledge of the spectral emissivity of the site is necessary for the application of the method. It should be applied over long time periods at each site to analyze the relationship between δT and $\delta(T_{31}-T_{32})$ and select the cases with $\delta(T_{31}-T_{32})$ values around zero for which the error introduced by the atmospheric profiles is small. In this study, the R-based method was validated using the ground LST measurements in the Valencia site and applied to all cases of the Valencia and Hainich sites (sections 4.2 and 4.3).

3. Experimental Data

3.1. Valencia Test Site

[17] The test site is located in a large (>30 km^2) area of rice fields south of Valencia, in the eastern Mediterranean coast of Spain. In July and August, rice crops attain nearly full vegetation cover and are well irrigated, which makes the site highly homogeneous in terms of both surface temperature and emissivity, thus easing the radiometric measurement of LST. Ground data from the Valencia test site have been used in previous studies [*Coll et al.*, 2005, 2006, 2007, 2009]. The thermal homogeneity of the site at daytime was assessed at 1×1 km^2 scale using MODIS data, and at 90×90 m^2 scale with ASTER TIR data. For 40 boxes of 11×11 ASTER pixels covering the rice field area, the standard deviation of the surface temperatures ranged between 0.2 and 0.7 K, with values lower than 0.5 K in most cases. The thermal heterogeneity was smaller at the MODIS scale, with typical standard deviations of 0.2–0.3 K for arrays of 3×3 pixels along the area.

3.1.1. In Situ LST Measurements

[18] In situ temperatures were measured at the Valencia site by several TIR radiometers distributed over a 1 km^2 grid within the rice field area concurrently with Terra/MODIS daytime overpasses (~1100 UTC). The instruments were two CIMEL CE 312–1 radiometers with four bands (8–13 μm, 11.5–12.5 μm, 10.5–11.5 μm and 8.2–9.2 μm) [*Sicard et al.*, 1999], one CIMEL CE 312–2 radiometer with six bands (8–13 μm, 8.1–8.5 μm, 8.5–8.9 μm, 8.9–9.3 μm, 10.3–11.0 μm, and 11.0–11.7 μm), one Apogee IRTS radiometer (single band, 6.5–14 μm) [*Bugbee et al.*, 1998], and two Everest model 112.2L thermometers (single band, 8–13 μm) (http://www.everestinterscience.com). The instruments were calibrated against a reference blackbody before and after each field measurement and intercompared in the field.

[19] Radiometers were placed about 150 m apart in the rice fields and carried back and forth along transects of 100–200 m in length. The in situ LSTs were calculated by averaging the ground temperatures measured by the available radiometers within three minutes centered at the satellite overpass time. The standard deviation of the ground temperatures was calculated as a measure of the spatial and temporal variability of LST in the test site (typically ≤ 0.5 K). More details on the ground LST measurements are given by *Coll et al.* [2005, 2006]. Radiometric temperatures were corrected for emissivity effects, including the reflection of the sky irradiance. Surface emissivity was measured in the field using the box method [*Rubio et al.*, 2003] for the four channels of the CE 312–1 radiometers. Details on the field emissivity measurements are given by *Coll et al.* [2007]. We obtained high emissivity ($\varepsilon = 0.985 \pm 0.005$) with negligible spectral variation in the 8–13 μm range, as expected for full vegetation cover.

[20] Together with the average in situ LST, we estimated the total uncertainty including the radiometer calibration error, the emissivity correction error (~0.2 K for measured emissivity uncertainty of 0.005), and the LST variability. A total of 23 cloud-free, daytime concurrences of ground and Terra/MODIS data were collected in July and August 2002–2007. Table 1 lists the in situ LSTs ($T_{in\ situ}$) and uncertainties for each case. The center of the 1 km^2 grid was at 0°17′50″W, 39°14′27″N for cases 1–3; 0°17′43″W, 39°15′01″N for cases 4–7; and 0°18′28″W, 39°15′54″N for cases 8–23.

3.1.2. MODIS Data

[21] Terra/MODIS data were acquired over the Valencia test site for all cases of Table 1. We used (1) brightness temperatures in bands 31 and 32 (T_{31} and T_{32}) from

Table 1. In Situ Measured LSTs and Uncertainties in the Valencia Test Site and Concurrent Terra/MODIS Data[a]

Case	Date	$T_{in\ situ} \pm \sigma$ (°C)	N	Granule ID	θ (deg)	T_{31} (°C)	T_{32} (°C)	T_{MOD11} (°C)	δT (K)
1	10 Jul 2002	28.8 ± 0.7	2	A2002191.1030	43.7	23.9	23.0	27.6	−1.2
2	11 Jul 2003	28.9 ± 0.8	3	A2003192.1040	27.7	26.7	26.3	29.5	0.6
3	12 Aug 2003	31.2 ± 0.6	4	A2003224.1040	28.1	28.3	27.8	31.2	0.0
4	8 Jul 2004	25.3 ± 0.6	2	A2004190.1020	50.3	22.5	22.0	25.4	0.1
5	27 Jul 2004	27.9 ± 0.6	2	A2004209.1050	5.6	25.5	24.9	28.3	0.4
6[b]	3 Aug 2004	30.0 ± 0.7	2	A2004216.1100	6.0	26.6	25.8	29.9	−0.1
7[b]	12 Aug 2004	28.7 ± 0.5	2	A2004225.1050	5.7	25.8	25.2	28.8	0.1
8	12 Jul 2005	27.2 ± 0.6	4	A2005193.1105	16.8	24.9	24.5	27.4	0.2
9	14 Jul 2005	27.9 ± 0.7	3	A2005195.1050	6.1	24.9	24.3	27.8	−0.1
10[b]	21 Jul 2005	28.4 ± 0.7	4	A2005202.1100	5.7	25.8	25.1	28.7	0.3
11	28 Jul 2005	28.9 ± 0.4	3	A2005209.1105	16.9	24.7	23.9	28.0	−0.9
12[b]	6 Aug 2005	28.3 ± 0.4	2	A2005218.1100	5.7	25.4	24.9	28.3	0.0
13	3 Jul 2006	29.9 ± 0.9	3	A2006184.1040	27.6	27.6	27.1	30.4	0.5
14	17 Jul 2006	29.9 ± 0.7	4	A2006198.1050	6.0	25.1	23.8	29.4	−0.5
15	22 Jul 2006	29.4 ± 0.8	3	A2006203.1110	27.1	26.2	25.5	29.4	0.0
16[b]	24 Jul 2006	29.2 ± 0.9	3	A2006205.1100	5.7	25.8	24.9	29.2	0.0
17	28 Jul 2006	28.5 ± 0.7	3	A2006209.1035	36.5	24.1	23.4	27.3	−1.2
18[b]	2 Aug 2006	29.7 ± 0.7	4	A2006214.1050	6.0	25.0	23.8	28.9	−0.8
19	3 Jul 2007	26.7 ± 0.9	5	A2007184.1010	60.3	21.4	20.1	26.3	−0.4
20	4 Jul 2007	28.8 ± 0.9	6	A2007185.1050	5.9	23.3	21.8	27.4	−1.4
21[b]	11 Jul 2007	27.1 ± 0.4	5	A2007192.1100	5.8	22.5	21.4	26.4	−0.7
22[b]	20 Jul 2007	28.1 ± 0.4	6	A2007201.1055	5.8	24.1	23.2	27.8	−0.3
23	26 Jul 2007	27.6 ± 0.4	6	A2007207.1015	55.6	22.7	21.9	26.3	−1.3

[a] $T_{in\ situ}$ indicates in situ measured LSTs, and σ indicates uncertainties. N is the number of ground radiometers used, θ is the satellite zenith angle, and T_{31} and T_{32} are the brightness temperatures in bands 31 and 32, respectively. T_{MOD11} is the MOD11_L2 LST, and $\delta T = T_{MOD11} - T_{in\ situ}$ is the LST error for the T-based validation.

[b] Local radiosonde launched around the overpass time.

MOD02 1 km calibrated and geolocated radiances, (2) satellite viewing zenith angle (θ) from the MOD03 product, (3) LST (T_{MOD11}) and surface emissivity in bands 31 and 32 (ε_{31} and ε_{32}) from the MOD11_L2 product, and (4) atmospheric column water vapor (W_{MOD}) and air temperature at surface level (T_{air}) data from the MOD07 product [*Seemann et al.*, 2006]. We extracted the values for the four pixels closest to the center of the test site and obtained the magnitude corresponding to the coordinates of the site by linear interpolation. Table 1 lists the MODIS data (θ, T_{31}, T_{32}, and T_{MOD11}) concurrent to the ground LST measurements. ε_{31} ranged from 0.982 to 0.984, and ε_{32} from 0.982 to 0.987, in good agreement with the field measurements. Although the nominal spatial resolution is 1 km^2, MODIS TIR pixels are actually sensitive to an area of 1 km (along-track) by 2 km (across-track) at nadir [*Barnes et al.*, 1998]. The MODIS footprint is larger for off-nadir pixels (e.g., 1.7 km × 6.6 km for $\theta = 60°$). Therefore, cases 19 and 23 with large viewing angles should be considered with care in the T-based validation.

[22] In order to increase the number of validation cases for the R-based method, we acquired an additional set of 51 Terra scenes in July–August 2002–2007. Table 2 lists the 51 additional validation cases. Cases 24–57 are daytime and cases 58–74 are nighttime (note that case number continues from Table 1). For daytime cases W_{MOD} ranged between 1.7 and 3.8 cm, and T_{MOD11} between 25 and 31°C. For nighttime cases W_{MOD} ranged from 2.5 to 4.5 cm and T_{MOD11} from 20 to 24°C.

3.1.3. Atmospheric Profiles From Local Radiosondes and NCEP Data

[23] Atmospheric profiles of pressure, temperature and humidity were measured at the Valencia test site by means of Vaisala RS80 radiosondes concurrent to Terra overpasses on eight cases (marked with a footnote in Table 1). Balloons were launched half an hour before the satellite overpass and reached 10 km of height in about 1 h. The atmospheric profiles were used as inputs to the MODTRAN 4 radiative transfer code [*Berk et al.*, 1999] for the application of the R-based method. The measured radiosonde profiles were completed with midlatitude summer standard profiles, including the fixed gases, up to 100 km altitude. A rural aerosol model with visibility of 23 km was selected. Spectral atmospheric transmittance and path radiance were calculated at the satellite viewing angle for each case, and sky downwelling radiance was obtained at several zenith angles to calculate the sky flux by integration over the hemisphere.

[24] Since radiosonde measurements are seldom and limited to dedicated validation campaigns, we also used the NCEP global tropospheric analyses product [*Kalnay et al.*, 1996], which provides global atmospheric data at 1° × 1° grids every 6 h. In the present work, we used geopotential height, temperature, and relative humidity at 26 standard pressure levels for the grid centered at 0°E, 39°N for all cases of Tables 1 and 2 (at 1200/0000 UTC for daytime/nighttime). NCEP data were downloaded from http://dss.ucar.edu/datasets/ds083.2/. As for the local profiles, the NCEP profiles were completed with standard profiles of fixed gases and rural aerosol model with 23 km of visibility.

3.2. Hainich Test Site

[25] The Hainich, Germany test site (51°04′45.14″N, 10°27′07.83″E, elevation of 445 m) is one of the CarboEurope-IP measurement sites (http://www.carboeurope.org). The land cover is mixed broadleaf deciduous forest dominated by European Beech (*Fagus sylvatica*). In the present study, the site was used for R-based validation only. Ground LST data are available at the Hainich site within the CarboEurope measurement program, and they were recently used by *Wang et al.* [2008] for the T-based validation of the

Table 2. Additional Terra/MODIS Data for the R-Based Validation in the Valencia Test Site[a]

Case	Granule ID	Date	θ (deg)	T_{MOD11} (°C)
24	A2002196.1045	15 Jul 2002	16.5	26.0
25	A2002203.1050	22 Jul 2002	5.0	23.9
26	A2002210.1100	29 Jul 2002	6.6	24.1
27	A2002219.1050	7 Aug 2002	5.7	27.3
28	A2002224.1110	12 Aug 2002	27.4	27.3
29	A2002226.1100	14 Aug 2002	6.3	26.4
30	A2003195.1110	14 Jul 2003	26.9	27.3
31	A2003197.1100	16 Jul 2003	5.4	28.3
32	A2003199.1045	18 Jul 2003	17.7	28.7
33	A2004191.1105	9 Jul 2004	17.0	25.3
34	A2004193.1050	11 Jul 2004	5.7	25.6
35	A2004198.1110	16 Jul 2004	27.3	26.0
36	A2004200.1100	18 Jul 2004	6.0	26.9
37	A2004202.1045	20 Jul 2004	17.0	27.5
38	A2004207.1105	25 Jul 2004	17.2	27.6
39	A2005184.1110	3 Jul 2005	27.1	28.4
40	A2005188.1045	7 Jul 2005	17.3	26.7
41	A2005223.1115	11 Aug 2005	36.0	27.2
42	A2005225.1105	13 Aug 2005	16.9	28.3
43	A2005227.1050	15 Aug 2005	6.0	27.9
44	A2006187.1110	6 Jul 2006	27.0	25.6
45	A2006191.1045	10 Jul 2006	17.5	29.0
46	A2006196.1105	15 Jul 2006	16.9	28.9
47	A2006219.1110	7 Aug 2006	27.0	27.6
48	A2006221.1100	9 Aug 2006	5.6	27.6
49	A2006223.1045	11 Aug 2006	17.4	28.3
50	A2006226.1115	14 Aug 2006	35.8	25.6
51	A2007208.1100	27 Jul 2007	6.0	28.0
52	A2007210.1045	29 Jul 2007	16.9	28.5
53	A2007213.1115	1 Aug 2007	36.1	26.3
54	A2007215.1105	3 Aug 2007	17.3	24.3
55	A2007217.1050	5 Aug 2007	5.5	26.9
56	A2007224.1100	12 Aug 2007	6.2	27.9
57	A2007226.1045	14 Aug 2007	16.9	27.2
58	A2003199.2150	18 Jul 2003	20.6	24.1
59	A2004200.2205	18 Jul 2004	2.9	22.7
60	A2004205.2220	23 Jul 2004	34.2	22.2
61	A2004207.2210	25 Jul 2004	14.3	24.1
62	A2005184.2215	3 Jul 2005	24.6	22.7
63	A2005186.2205	5 Jul 2005	2.5	22.4
64	A2005227.2200	15 Aug 2005	8.9	22.2
65	A2006189.2205	8 Jul 2006	2.7	23.9
66	A2006191.2150	10 Jul 2006	19.8	22.8
67	A2006196.2210	15 Jul 2006	14.6	23.3
68	A2006221.2205	9 Aug 2006	2.6	23.3
69	A2006223.2154	11 Aug 2006	19.9	22.7
70	A2007208.2205	27 Jul 2007	3.1	23.7
71	A2007210.2150	29 Jul 2007	19.5	22.5
72	A2007215.2210	3 Aug 2007	14.7	23.5
73	A2007222.2220	10 Aug 2007	25.3	22.6
74	A2007224.2205	12 Aug 2007	3.1	19.6

[a]Case number continues from Table 1. θ is the satellite zenith angle, and T_{MOD11} is the MOD11_L2 LST.

MOD11_L2 LST product in nighttime conditions. At the Hainich site, ground LSTs correspond to a single point covering a small area of a few meters in diameter above the forest canopy and are derived from upwelling longwave radiation measurements from a tower at 44 m above ground [*Wang et al.*, 2008]. Single point measurements may not be representative of LSTs in MODIS 1 km pixels especially for such a complex forest surface, even at nighttime. Another major error source in the ground LSTs estimated from longwave radiation measurements is that the atmospheric effects (due to air temperatures slightly warmer than canopy temperatures at night) in the layers between the instrument and the canopy at different heights in the entire hemisphere is not considered [see *Wang et al.*, 2008, equation 2].

[26] We recognize that it is very difficult to make ground LST measurements over a forest at the MODIS scale. Therefore, the R-based validation method may be more feasible in this case. The Hainich site was selected in this study because of the availability of atmospheric profile measurements from a nearby radio sounding station in Meiningen, Germany (50.56°N, 10.38°E, elevation of 453 m; about 60 km south of Hainich). To avoid extrapolation, the atmospheric profiles at and above 453 m were used in the radiative transfer simulations at the Hainich site. The surface spectral emissivity of grassland measured in the Texas field campaign in 2005 [*Wan and Li*, 2008] was used here because the values (ε_{31} = 0.979 and ε_{32} = 0.982) are appropriate for vegetated surfaces and are similar to the constant broadband emissivity (0.980) used by *Wang et al.* [2008]. The emissivity values are also consistent with field measurements of beech canopies reported by *Ribeiro da Luz and Crowley* [2007], and with the results of *Momeni and Saradjian* [2007] for partially and fully vegetated canopies.

[27] The emissivity assumption was assessed using MOD13A2 vegetation indices over the Hainich forest along 2004 and the Texas grassland site where the emissivity spectra were measured. MOD13A2 products provide 1 km, 16-day Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI), which can be related to emissivity. Figure 1 shows the MOD13A2 NDVI and EVI values for the two sites. At the time of the emissivity measurements (21 April), we obtained NDVI = 0.36 and EVI = 0.20 at the grassland site. According to Figure 1a, 19 out of 23 NDVI values range from 0.4 to 0.9 at the Hainich

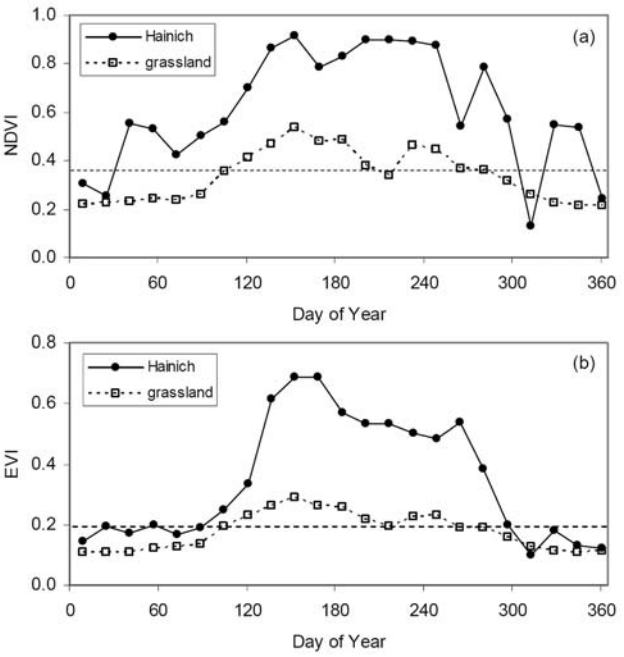

Figure 1. Time series of MOD13A2 vegetation indexes over the Hainich site and the Texas grassland site. The horizontal dashed line shows the grassland vegetation index at the time of the emissivity measurements (21 April). (a) NDVI. (b) EVI.

Table 3. Validation Cases for the Hainich Test Site in March–October 2004[a]

Case	Granule ID	Date	θ (deg)	T_{M*D11} (°C)	W_{rad} (cm)	W_{NCEP} (cm)
1	A2004077.0935	17 Mar	56.9	12.2	0.74	1.07
2	A2004077.2045	17 Mar	26.3	9.6	1.22	1.51
3	A2004107.0945	16 Apr	48.3	19.4	1.02	0.95
4	A2004107.2100	16 Apr	9.8	7.6	0.89	1.00
5	A2004135.2125	14 May	26.5	6.7	1.08	1.39
6	A2004137.1000	16 May	36.1	15.4	1.03	1.52
7	A2004166.2040	14 Jun	33.5	13.3	2.28	2.37
8	A2004167.1010	15 Jun	21.3	20.6	2.00	2.21
9	A2004198.2040	16 Jul	33.4	15.4	2.20	2.30
10	A2004199.1010	17 Jul	21.2	23.1	2.62	2.72
11	A2004228.1120	15 Aug	56.5	17.7	2.00	1.72
12	A2004228.2050	15 Aug	18.2	14.0	1.67	2.48
13	A2004258.0955	14 Sep	42.5	16.0	2.23	1.40
14	A2004262.2040	18 Sep	33.3	13.2	1.22	2.54
15	A2004292.0940	18 Oct	52.3	7.4	1.10	1.41
16	A2004292.2050	18 Oct	18.0	5.6	0.78	0.93
17	A2004303.1100	29 Oct	43.4	13.9	1.41	1.99
18	A2004303.2035	29 Oct	39.6	6.6	1.72	1.51
19	A2004082.0955	22 Mar	42.7	6.2	0.74	0.91
20	A2004088.2030	28 Mar	45.5	0.9	0.53	1.01
21	A2004101.1025	10 Apr	3.9	11.7	0.72	0.78
22	A2004102.2040	11 Apr	33.9	0.2	0.64	1.01
23	A2004144.2115	23 May	18.3	4.1	0.66	0.92
24	A2004147.1035	26 May	15.2	11.6	1.16	0.96
25	A2004158.2130	6 Jun	34.0	12.1	1.84	2.06
26	A2004176.1005	24 Jun	29.4	14.4	1.05	1.23
27	A2004203.0945	21 Jul	48.0	20.7	2.45	2.01
28	A2004206.2130	24 Jul	34.1	12.7	1.24	2.02
29	A2004222.1020	9 Aug	12.3	25.1	1.82	1.63
30	A2004222.2130	9 Aug	34.0	16.7	1.59	1.84
31	A2004249.2110	5 Sep	9.3	14.9	1.51	1.40
32	A2004250.1040	6 Sep	23.7	21.7	0.97	1.66
33	A2004279.1010	5 Oct	20.8	19.2	2.11	2.41
34	A2004279.2120	5 Oct	26.8	14.5	2.25	2.64
35	A2004077.0145	17 Mar	23.4	5.4	0.70	1.05
36	A2004077.1120	17 Mar	54.0	17.7	0.74	1.07
37	A2004106.0115	15 Apr	21.8	2.9	0.65	0.75
38	A2004106.1225	15 Apr	26.4	20.5	0.78	0.78
39	A2004140.0105	19 May	36.2	9.9	1.39	0.89
40	A2004140.1215	19 May	9.8	20.0	1.47	1.53
41	A2004176.0215	24 Jun	52.8	8.8	1.32	1.72
42	A2004176.1150	24 Jun	25.8	17.3	1.05	1.23
43	A2004196.0155	14 Jul	31.1	7.7	1.58	1.21
44	A2004196.1125	14 Jul	49.7	15.8	1.32	2.04
45	A2004225.0120	12 Aug	12.8	17.0	2.09	1.91
46	A2004228.1125	15 Aug	49.9	21.1	2.00	1.72
47	A2004261.0055	17 Sep	42.9	7.1	0.73	1.33
48	A2004261.1210	17 Sep	0.4	16.1	0.80	1.10
49	A2004296.0130	22 Oct	3.3	6.8	0.64	1.29
50	A2004296.1240	22 Oct	40.3	14.0	0.81	1.34

[a]The sixth and seventh columns show the column water vapor from radiosonde profiles and NCEP data, respectively. Cases 1–34 are from Terra. Cases 35–50 are from Aqua.

site. Figure 1b shows that the EVI values at Hainich are close to or larger than 0.2 for most of the year. For NDVI values larger than 0.3, emissivities at 11 and 12 μm show negligible variation with NDVI [*French et al.*, 2008]. This means that, during most of the time, the real surface emissivity at the Hainich site may be similar to the values measured at the Texas grassland site. In order to avoid possible emissivity variations due to snow cover (as suggested by some low NDVI values in Figure 1a) and senescence in the nongrowing season, we only used data from March to October. The error related to possible uncertainty in surface emissivity is analyzed in sections 4.2 and 4.3.

[28] We selected 34 Terra and 16 Aqua cloud-free scenes over the Hainich site from March to October 2004, including daytime and nighttime (Table 3). We extracted the values of the four pixels closest to the test site for the brightness temperatures in bands 31 and 32 (M*D02 1 km), the satellite viewing zenith angle (M*D03), and T_{M*D11} (M*D11_L2). As before, the values at the site were obtained by linear interpolation. In Table 3, MODIS LSTs showed a large range (0–25°C,) owing to diurnal and seasonal variations.

[29] The radiosonde data measured at Meiningen corresponding to each of the 50 MODIS scenes (at 0000/1200 UTC for nighttime/daytime cases) were downloaded from http://weather.uwyo.edu/upperair/sounding.html thanks to the Department of Atmospheric Science, University of Wyoming. Water vapor content derived from the radiosonde profiles (W_{rad}, see Table 3) shows a range between 0.5 and 2.6 cm. As an alternative source of atmospheric profiles, we also used NCEP data at the Hainich site. For the NCEP grid closest to Hainich, the elevation of the surface level varies mostly between 50 and 250 m. We interpolated the two lower levels of the NCEP profiles to obtain the surface level at 445 m for the Hainich site. Column water vapor derived from the NCEP data (W_{NCEP}) are shown in Table 3.

4. Results and Discussion

4.1. T-Based Validation: Valencia Test Site

[30] The tenth column of Table 1 gives the LST errors $\delta T = T_{MOD11} - T_{in\ situ}$ for each validation case. Overall, the average value of δT (or bias) is -0.29 K, with standard deviation of 0.60 K, and root mean square error (RMSE) of ±0.67 K. δT ranges from -1.4 to 0.6 K and is within ±1.0 K for 19 cases. The difference between T_{MOD11} and T_{31}, which represents the magnitude of the atmospheric and emissivity correction, ranged between 2.5 and 4.9 K for the 23 cases. The variability is relatively large, given that the data used in this work cover rather limited regimes in terms of surface temperature and atmospheric conditions.

[31] Although the number of validation cases is small, we found certain correlation of δT with the brightness temperature difference, $T_{31} - T_{32}$ (Figure 2a) and the path water vapor content, $W_{MOD}/\cos\theta$ (Figure 2b). For the data shown here, the largest negative values of δT occurred for large values of both $T_{31} - T_{32}$ and atmospheric water vapor (e.g., case 20 in Table 1, with $\delta T = -1.4$ K, $T_{31} - T_{32} = 1.5$ K and $W_{MOD} = 3.8$ cm). It appears that the MOD11_L2 product underestimated the LSTs in the cases where the atmospheric correction was expected to be larger. On the other hand, a low correlation ($R^2 = 0.23$) was obtained between δT and the difference between the in situ LST and the air temperature (not shown), possibly because the effect of T_{air} is taken into account by the generalized split window algorithm. No significant correlation was found with the in situ LST, the MODIS LST and the viewing angle. However, results cannot be conclusive since the data set covers only narrow ranges of these variables and most of the viewing angles are close to nadir.

[32] In Figure 2a, four cases (1, 11, 17 and 23 in Table 1) lie outside the observed trend, with $T_{31} - T_{32}$ values of 0.7–0.9 K and anomalous negative values of δT between -0.9

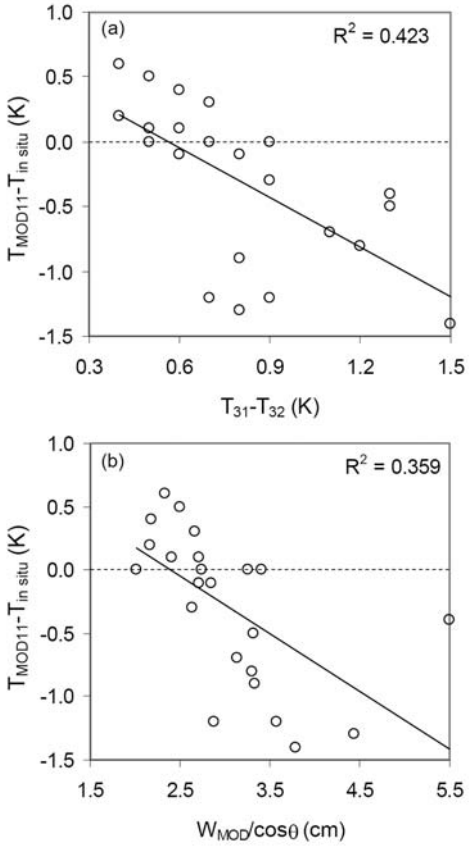

Figure 2. MOD11_L2 LST minus in situ LST for the T-based validation in the Valencia test site against (a) brightness temperature difference, $T_{31}-T_{32}$, and (b) MOD07 column water vapor, W_{MOD}. The solid line is the linear regression, and R^2 is the coefficient of determination.

and −1.3 K. These points can be also identified as some of the outliers in Figure 2b. The atypical behavior may be attributed to undetected cirrus cloud contamination or heavy aerosol loadings. Moreover, case 23 corresponds to a very large viewing angle (55°). If the four cases were excluded, the overall bias would reduce to −0.11 K and the standard deviation to 0.49 K (RMSE = ±0.50 K).

4.2. R-Based Validation: Valencia Test Site

[33] In the R-based method, the in situ LSTs were calculated from the brightness temperatures in band 31 using MODTRAN 4 radiative transfer simulations based on locally measured atmospheric profiles for eight cases in Table 1, and NCEP profiles for all the 74 cases in Tables 1 and 2. The difference between the product LST and the calculated LST ($T_{R\text{-}based}$) is the estimated accuracy of the product. For each validation case, we obtained the difference $\delta(T_{31}-T_{32})$ between the actual MODIS and the profile-based calculated $T_{31}-T_{32}$ values. Emissivities in bands 31 and 32 were obtained from the MOD11_L2 product. Tables 4 and 5 show the values of $T_{R\text{-}based}$, $\delta T = T_{MOD11} - T_{R\text{-}based}$, and $\delta(T_{31}-T_{32})$ obtained from the local radiosonde profiles and the NCEP data, respectively. Column water vapor and atmospheric transmittance in band 31 (τ_{31}) are also given. First, the accuracy of $T_{R\text{-}based}$ is assessed by means of the concurrent ground LST measurements and an uncertainty analysis is presented. Then, the R-based validation of the MOD11_L2 product is discussed.

4.2.1. Uncertainty Analysis of the R-Based Method

[34] The accuracy of the R-based method was assessed with a comparison between the ground LST data in Table 1 and the corresponding $T_{R\text{-}based}$ calculated for the eight cases with local radiosonde profiles (Table 4) and the 23 cases with NCEP profiles (cases 1–23 in Table 5). Results are shown in Figure 3, where the difference between the ground-measured LST and the R-based LST ($\delta T = T_{in\ situ} - T_{R\text{-}based}$) is plotted against $\delta(T_{31}-T_{32})$. Figure 3 shows a good correlation between δT and $\delta(T_{31}-T_{32})$ ($R^2 = 0.67$), with the smallest LST errors corresponding to a narrow range of $\delta(T_{31}-T_{32})$ values in the vicinity of zero. For most cases with radiosonde profiles, δT ranged between −0.1 and 0.8 K, and $\delta(T_{31}-T_{32})$ between 0.0 and 0.4 K, except for case 18 (δT = 1.5 K and $\delta(T_{31}-T_{32})$ = 0.5 K). As expected, NCEP profiles yielded larger δT and $\delta(T_{31}-T_{32})$ ranges. Large negative δT values (−1.4 to −1.9 K in cases 1, 14 and 18) correspond to large negative $\delta(T_{31}-T_{32})$ values (−0.5 to −0.7 K), showing that NCEP overestimated the atmospheric effect.

[35] The comparison between ground and R-based LSTs in Figure 3 is affected by uncertainties in both the ground LSTs (section 3.1.1) and the radiative transfer calculations of the in situ LST and $\delta(T_{31}-T_{32})$. An important part of the error may be due to inaccuracies in the temperature and water vapor profiles or spatial and temporal atmospheric variations. A balloon typically takes 1 h to reach the elevation level of 10 km, and thus to measure the entire water vapor profile. Therefore, radiosonde measurements may not represent the real atmospheric conditions at the times of instantaneous satellite observations. On the other hand, uncertainties in the assumed emissivity have an impact on the LST error.

[36] We estimated the uncertainty in the calculation of $T_{R\text{-}based}$ by means of a simulation study using case 6 in Table 4, which is an average case in terms of column water vapor (W_{rad} = 2.4 cm). The effect of atmospheric variations was simulated in two ways. First, the water vapor mixing ratio was increased by 10% at each profile level. Second, the air temperature was increased by 1 K at each level from surface up to 5 km. These changes can be regarded as small temporal and spatial atmospheric variations or typical errors in radiosondes profiles. For each

Table 4. R-Based Validation Results for the Valencia Test Site With Local Radiosonde Profiles[a]

Case	W_{rad} (cm)	τ_{31}	$T_{R\text{-}based}$ (°C)	δT (K)	$\delta(T_{31}-T_{32})$ (K)
6	2.37	0.777	29.8	0.1	0.12
7	2.09	0.769	28.4	0.4	0.17
10	2.01	0.763	28.2	0.5	0.44
12	1.86	0.824	27.5	0.8	0.30
16	2.39	0.714	28.5	0.7	0.34
18	2.86	0.669	28.2	0.7	0.46
21	2.90	0.643	27.2	−0.8	0.01
22	2.90	0.615	28.2	−0.4	0.00

[a] W_{rad} is the column water vapor calculated from the profiles. $\delta T = T_{MOD11} - T_{R\text{-}based}$ (the corresponding T_{MOD11} data are shown in Table 1). $\delta(T_{31}-T_{32})$ is the difference between the MODIS (actual) and the profile-based (calculated) brightness temperature difference in bands 31 and 32.

Table 5. R-Based Validation Results for the Valencia Test Site With NCEP Atmospheric Profiles[a]

Case	W_{NCEP} (cm)	τ_{31}	$T_{R\text{-based}}$ (K)	δT (K)	$\delta(T_{31}-T_{32})$ (K)
1	2.97	0.591	30.5	−2.9	−0.72 F
2	1.80	0.833	29.2	0.3	0.05
3	1.26	0.891	30.7	0.5	0.28
4	1.08	0.872	25.6	−0.2	0.01
5	1.50	0.875	28.0	0.3	0.38
6	1.82	0.833	29.9	0.0	0.22
7	2.32	0.795	29.1	−0.3	−0.06
8	2.14	0.794	28.0	−0.6	−0.14
9	2.53	0.753	27.3	0.5	0.28
10	2.34	0.789	28.7	0.0	0.22
11	3.20	0.694	29.6	−1.6	−0.48 F
12	2.55	0.759	28.2	0.1	0.03
13	1.56	0.867	30.6	−0.2	0.07
14	3.79	0.584	31.8	−2.4	−0.46 F
15	2.16	0.806	29.4	0.0	0.13
16	3.07	0.696	30.0	−0.8	−0.11
17	2.04	0.781	27.7	−0.4	0.04
18	3.94	0.587	31.5	−2.6	−0.55 F
19	2.56	0.607	26.5	−0.2	0.11
20	3.32	0.643	28.6	−1.2	0.16
21	2.50	0.752	26.5	−0.1	0.16
22	1.71	0.866	26.9	0.9	0.42
23	2.17	0.685	25.4	0.9	0.32
24	2.2	0.742	28.3	−2.3	−0.76 F
25	1.9	0.833	25.5	−1.6	−0.74 F
26	2.6	0.747	24.9	−0.8	−0.27
27	2.7	0.721	27.8	−0.5	−0.01
28	2.9	0.667	28.8	−1.5	−0.09
29	3.1	0.725	26.9	−0.5	−0.18
30	3.3	0.650	29.5	−2.3	−0.70 F
31	1.7	0.848	29.2	−0.8	−0.20
32	1.3	0.896	28.8	−0.1	−0.02
33	1.7	0.854	25.6	−0.3	0.07
34	2.6	0.726	24.8	0.8	0.45
35	3.7	0.594	29.8	−3.8	−1.13 F
36	3.2	0.686	28.5	−1.5	−0.40 F
37	2.9	0.736	28.0	−0.6	0.04
38	2.8	0.740	29.3	−1.7	−0.51 F
39	2.5	0.745	29.3	−0.9	−0.26
40	3.3	0.620	28.4	−1.6	−0.33 F
41	3.4	0.582	30.4	−3.1	−0.76 F
42	2.2	0.819	29.4	−1.0	−0.19
43	2.5	0.761	28.1	−0.2	0.09
44	2.8	0.678	25.9	−0.3	0.14
45	2.5	0.758	29.0	0.0	0.11
46	3.3	0.680	31.7	−2.8	−0.89 F
47	2.1	0.762	26.9	0.7	0.53 F
48	2.0	0.816	26.3	1.3	0.80 F
49	1.6	0.853	26.9	1.4	0.82 F
50	2.9	0.650	27.6	−1.9	−0.35 F
51	1.7	0.861	26.9	1.2	0.64 F
52	2.1	0.812	27.8	0.7	0.41
53	1.7	0.856	26.9	−0.7	0.08
54	2.1	0.798	22.4	1.9	0.67 F
55	2.9	0.759	28.9	−2.0	−0.63 F
56	2.4	0.791	28.8	−0.8	−0.23
57	2.5	0.810	27.8	−0.6	−0.02
58	1.3	0.894	24.0	0.1	0.06
59	3.1	0.672	23.9	−1.1	−0.33 F
60	3.7	0.594	24.5	−2.3	−0.39 F
61	2.7	0.694	25.5	−1.4	−0.17
62	2.8	0.722	23.9	−1.1	−0.17
63	2.3	0.746	22.3	0.1	0.27
64	2.4	0.778	22.3	0.0	0.05
65	3.2	0.655	23.9	0.0	0.04
66	2.6	0.744	22.9	−0.1	0.05
67	3.6	0.633	26.4	−3.1	−0.92 F
68	2.5	0.741	22.8	0.5	0.14
69	1.7	0.838	23.0	−0.3	0.24
70	2.5	0.758	24.4	−0.7	−0.08
71	2.5	0.728	23.5	−1.0	0.24
72	2.3	0.761	22.6	0.9	0.58 F
73	1.7	0.838	21.9	0.7	0.57 F
74	2.8	0.687	20.7	−1.2	−0.16

[a] W_{NCEP} is the column water vapor calculated from NCEP data. $\delta T = T_{MOD11} - T_{R\text{-based}}$ (T_{MOD11} data are shown in Table 1 for cases 1–23, and in Table 2 for cases 24–74). $\delta(T_{31}-T_{32})$ is the difference between the MODIS (actual) and the profile-based (calculated) brightness temperature difference in bands 31 and 32. F indicates cases failing the $\delta(T_{31}-T_{32})$ test.

change, $T_{R\text{-based}}$ was calculated from T_{31} with the modified profile (inverse simulation) and compared with the value obtained with the original profile. The differences (in absolute value) were 0.35 K for the water vapor variation and 0.26 K for the air temperature variation. The total atmospheric effect on $T_{R\text{-based}}$ was calculated from the above values yielding RMSE = ±0.44 K. On the other hand, *Wan and Li* [2008] compared MODTRAN 4 simulations with clear-sky radiances measured by an advanced Bomem TIR interferometer (MR100) well calibrated by two blackbodies to accuracy better than 0.1 K for the brightness temperature of radiance.

[37] The impact of the surface emissivity was evaluated assuming an uncertainty of 0.005 in ε_{31}. This is justified since vegetated surfaces show high emissivities with small variations at this band, and it is a typical uncertainty in field emissivity measurements [*Wan*, 2008]. However, larger uncertainties or variations may be expected for bare surfaces. For the present case, the emissivity uncertainty resulted in an error of 0.25 K in the calculated LST. The total error in $T_{R\text{-based}}$ was obtained from the above atmospheric, radiative transfer model and emissivity uncertainties, yielding RMSE = ±0.5 K. The largest part of the error is due to the atmospheric variations. The uncertainty might be larger if the differences between the atmospheric profiles and the real atmospheric conditions exceed the values assumed here.

[38] The analysis was repeated with the radiosonde profiles of cases 26 and 27 in Table 3 (Hainich site) corresponding to drier and wetter conditions ($W_{rad}/\cos\theta$ of 1.2 and 3.7 cm, respectively). Using the same approach, the total error in the calculated $T_{R\text{-based}}$ was ±0.4 and ±0.8 K, with the atmospheric error increasing with the water vapor content. These error figures are similar to the uncertainties of the ground measured LSTs and are compatible with the δT differences for radiosonde profiles in Figure 3.

[39] The uncertainty in the calculation of the profile based $T_{31}-T_{32}$ was estimated using a similar approach. At-sensor brightness temperatures for the two MODIS bands were calculated in so-called forward simulations. For case 6 in Table 4, the effect of increasing water vapor by 10% was of −0.26 K in T_{31} and −0.38 K in T_{32}, resulting in a change of 0.12 K in $T_{31}-T_{32}$. The effect of increasing air temperatures by 1 K was 0.20 K in T_{31} and 0.30 K in T_{32}, resulting in a change of −0.10 K for $T_{31}-T_{32}$. The RMSE of the two atmospheric effects is 0.16 K. In some cases, the mismatches between the profiles used in the R-based calculations and the actual profiles along the path of MODIS

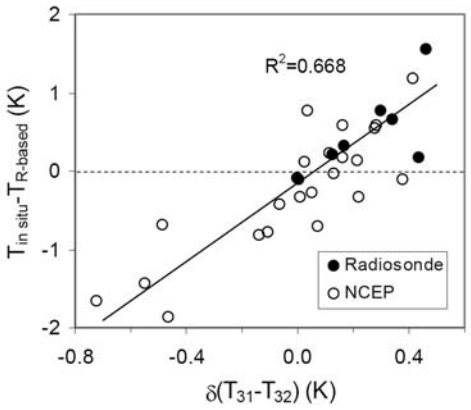

Figure 3. Ground-measured LST minus calculated LST with the R-based method from local radiosonde profiles (closed circles) and NCEP profiles (open circles) against $\delta(T_{31}-T_{32})$ in the Valencia site. The solid line is the linear regression, and R^2 is the coefficient of determination.

observations may be 2–4 times larger (20–40% difference in water vapor), resulting in changes in the $T_{31}-T_{32}$ values up to 0.64 K or larger. For an emissivity uncertainty of 0.005 in the two MODIS bands, we obtained variations of 0.19 K in T_{31} and 0.17 K in T_{32}. There may be some dynamic emissivity errors due to possible emissivity variation with viewing angles or seasonal changes. For fully vegetated or bare soil sites, the dynamic emissivity error has equal sign and most likely equal magnitude in bands 31 and 32, resulting in a negligible change in $T_{31}-T_{32}$. However, emissivity errors may be larger and may have different signs and magnitudes in sparsely vegetated areas.

[40] Local, near-concurrent radiosonde profiles likely provide the most reliable description of the atmospheric state for radiative transfer calculations. NCEP profiles may be less accurate than assumed in the previous uncertainty analysis, thus the errors in the calculated $T_{R\text{-based}}$ and $T_{31}-T_{32}$ may be larger. To assess the accuracy of the NCEP profiles, we compared the results for the eight cases with local radiosonde of Table 4 with the same cases with NCEP profiles of Table 5. For these cases, the difference between radiosonde and NCEP column water vapor values ranges from −1.1 (−41%, case 18) to 1.2 cm (+38%, case 22). If we assume that the local radiosondes represent accurately the atmospheric state, then the profile-based in situ LST and $T_{31}-T_{32}$ values are correct. Therefore, the difference between $T_{R\text{-based}}$ calculated from radiosonde and NCEP profiles is the LST error due to the NCEP profile. Similarly, the difference between the radiosonde $T_{31}-T_{32}$ and the corresponding NCEP $T_{31}-T_{32}$ value is the $\delta(T_{31}-T_{32})$ difference due to the incorrect NCEP profile.

[41] Figures 4a and 4b show the difference δT between radiosonde and NCEP $T_{R\text{-based}}$ values against the column water vapor difference $W_{rad}-W_{NCEP}$ and $\delta(T_{31}-T_{32})$, respectively. δT values range from −3.3 (case 18) to 1.3 K (case 22) and show a good correlation ($R^2 = 0.80$) with the water vapor differences, with NCEP overestimating (underestimating) radiosonde $T_{R\text{-based}}$ when NCEP water vapor overestimated (underestimated) radiosonde water vapor. Furthermore, the correlation between δT and $\delta(T_{31}-T_{32})$ is excellent ($R^2 = 0.98$) and shows a similar behavior as in Figure 3. The correlation in Figure 4a is lower because the difference in the R-based LSTs depends not only on the column water vapor difference but also on the difference in temperature profiles and the MODIS viewing angle. The correlation in Figure 4b is very high because all these effects are also largely reflected in $\delta(T_{31}-T_{32})$. Results of Figure 4b indicate that $\delta(T_{31}-T_{32})$ can be used to check the suitability of the atmospheric profiles used for the R-based validation in most clear-sky conditions. According to these data, one can infer that for obtaining $T_{R\text{-based}}$ within a desirable accuracy of ±1.0 K, $\delta(T_{31}-T_{32})$ should be within ±0.3 K. The error bounds are in agreement with the above uncertainty analysis.

[42] Using the data in Figure 3, the linear regression between $\delta T = T_{in\ situ} - T_{R\text{-based}}$ and $\delta(T_{31}-T_{32})$ yielded $\delta T = 2.4 \times \delta(T_{31}-T_{32}) - 0.2$. From this equation, we can set the condition $-0.3\ K < \delta(T_{31}-T_{32}) < 0.5\ K$ for which the LST errors are within ±1.0 K. These $\delta(T_{31}-T_{32})$ limits are less restrictive but still consistent with the previous results. Since they are based on a comparison with independent ground LST measurements and include all cases with local radiosondes, we can select the condition $-0.3\ K < \delta(T_{31}-T_{32}) < 0.5\ K$ as a quality check for the atmospheric profiles used in the R-based method. However, the $\delta(T_{31}-T_{32})$ limits are not symmetric with respect to zero. This could be due to the sensitivity of the calculated $\delta(T_{31}-T_{32})$ values to the emissivity difference in bands 31 and 32, $\Delta\varepsilon$ =

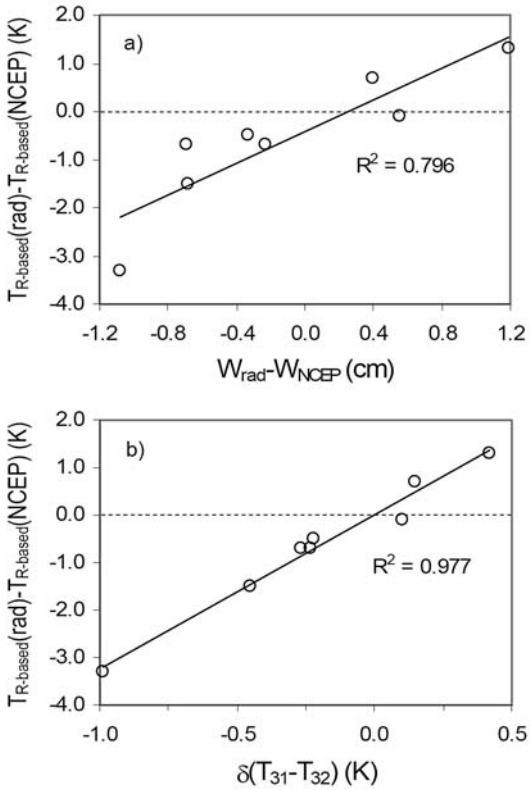

Figure 4. Difference between R-based LSTs obtained from local radiosonde and NCEP profiles in the Valencia site against (a) column water vapor difference and (b) difference in $T_{31}-T_{32}$ simulated from local radiosonde and NCEP profiles.

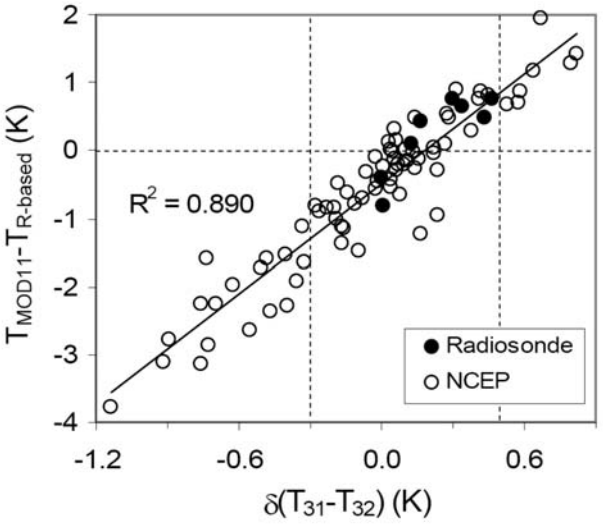

Figure 5. MOD11_L2 LST minus R-based in situ LST calculated with local radiosonde profiles and NCEP profiles against $\delta(T_{31}-T_{32})$ in the Valencia site. The solid line is the linear regression, and R^2 is the coefficient of determination. Vertical lines mark the valid $\delta(T_{31}-T_{32})$ range.

$\varepsilon_{31}-\varepsilon_{32}$. For the Valencia site, MOD11_L2 emissivities show an average value $\Delta\varepsilon = -0.003$. If $\Delta\varepsilon$ changes sign (from −0.003 to +0.003), the $\delta(T_{31}-T_{32})$ values in Tables 4 and 5, and Figure 3 would be changed by −0.2 K on average. Another reason could be the differential accuracy of the MODTRAN code in bands 31 and 32 due to uncertainties in the water vapor continuum absorption [*Dash and Ignatov*, 2008].

[43] The results shown here suggest that large errors in R-based LSTs are due to inappropriate atmospheric profiles, and that the $\delta(T_{31}-T_{32})$ test may be used as a quality check to discriminate the profiles used in the R-based method that reasonably represent the real atmosphere for the MODIS observation. According to the present data, it appears that cases meeting the condition $-0.3\ K < \delta(T_{31}-T_{32}) < 0.5\ K$ correspond to errors in the calculated in situ LST within ±1.0 K, which is appropriate for LST validation. Such LST errors are comparable with the uncertainty in the ground LST measurements and the typical LST errors for most cases of the T-based validation (Table 1). The R-based method and the $\delta(T_{31}-T_{32})$ test require that the surface emissivities in MODIS bands 31 and 32 are well known and show only small spatial and temporal variations, as in the case of vegetated areas.

4.2.2. R-Based Validation Results

[44] Figure 5 plots the MODIS LST errors ($\delta T = T_{MOD11} - T_{R\text{-}based}$) against the differences $\delta(T_{31}-T_{32})$ for the eight cases with local radiosonde profiles (Table 4) and the 74 cases with NCEP profiles (Table 5). In close agreement with Figure 3, Figure 5 shows a high correlation ($R^2 = 0.89$) between the R-based LST errors and $\delta(T_{31}-T_{32})$, with large positive (negative) values of δT corresponding to large positive (negative) values of $\delta(T_{31}-T_{32})$, and small absolute values of δT associated with small absolute values of $\delta(T_{31}-T_{32})$. Therefore, we can apply the condition $-0.3\ K < \delta(T_{31}-T_{32}) < 0.5\ K$ to select accurate atmospheric profiles for the R-based validation cases.

[45] All the cases with local radiosonde measurements are within the valid $\delta(T_{31}-T_{32})$ thresholds and δT ranges between −0.8 and 0.8 K. Cases 18 and 21 have almost the same W_{rad} values (2.9 cm) but the former underestimated and the latter overestimated the water vapor resulting in LST errors (0.7 and −0.8 K) and $\delta(T_{31}-T_{32})$ values (0.46 and 0.01 K) by their high and low ends. For the eight radiosonde cases, the MOD11_L2 product yielded average overestimation of 0.24 K, standard deviation of 0.58 K and RMSE = ±0.63 K.

[46] For the cases with NCEP profiles, the ranges of δT and $\delta(T_{31}-T_{32})$ were much larger due to inappropriate atmospheric profiles in several validation cases. In Table 5, cases failing to pass the above $\delta(T_{31}-T_{32})$ condition are marked with F. Considering only the cases passing the test (49 out of 74), we obtained average LST bias of −0.28 K, standard deviation of 0.47 K and RMSE = ±0.54 K, with δT ranging between −1.5 and 0.9 K (90% within ±1.0 K). The error figures are comparable to the T-based validation results (section 4.1). It suggests that the R-based method could be used for LST validation when the atmospheric profiles used in the radiative transfer simulations meet certain quality conditions in terms of $\delta(T_{31}-T_{32})$.

[47] We analyzed the dependence of the LST errors on the column water vapor (W_{NCEP}), and the atmospheric transmittance in band 31 (τ_{31}). We found that the larger negative LST errors occur for cases with higher W_{NCEP} and subsequently lower τ_{31}, that is, when radiative transfer calculations may produce larger uncertainties. Atmospheric profiles with large column water vapor (>3 cm) or low τ_{31} (<0.6) are usually associated with large, negative $\delta(T_{31}-T_{32})$ values. Therefore, the application of the R-based method for cases with large water vapor loads may be limited by the suitability of the atmospheric profiles.

[48] As an alternative to the $\delta(T_{31}-T_{32})$ threshold, we can go one step further and use the linear regression from Figure 3 ($T_{in\ situ} - T_{R\text{-}based} = 2.4 \times \delta(T_{31}-T_{32}) - 0.2$) to correct the $\delta(T_{31}-T_{32})$ dependence in the $T_{MOD11} - T_{R\text{-}based}$ differences in Figure 5 for the effect of the mismatches in the atmospheric profiles. We call this alternative as $\delta(T_{31}-T_{32})$ correction method, with which the R-based LST can be corrected to $T_{R\text{-}based}(cor) = T_{R\text{-}based} + 2.4 \times \delta(T_{31}-T_{32}) - 0.2$. The correction equation applied here is based on ground LST measurements and accurate emissivity values for fully vegetated surfaces. However, it is not a universal correction but depends on the range of atmospheric water vapor and effects of uncertainties in the assigned surface emissivities at a specific site. The difference between T_{MOD11} and $T_{R\text{-}based}(cor)$ is plotted against $\delta(T_{31}-T_{32})$ in Figure 6 for both radiosonde and NCEP profiles. The dependence of the LST errors on $\delta(T_{31}-T_{32})$ shown in Figure 5 disappeared. Therefore, we can take all validation cases regardless the $\delta(T_{31}-T_{32})$ value. For the corrected R-based LSTs in all 74 NCEP cases, the mean LST bias was −0.35 K, the standard deviation was 0.41 K and RMSE = ±0.54 K, in close agreement with the results for the cases passing the $\delta(T_{31}-T_{32})$ test but without using the correction method.

4.3. R-Based Validation: Hainich Test Site

[49] The Hainich forest validation data sets encompass both Terra and Aqua, daytime and nighttime scenes covering the seasonal variability of surface and atmospheric

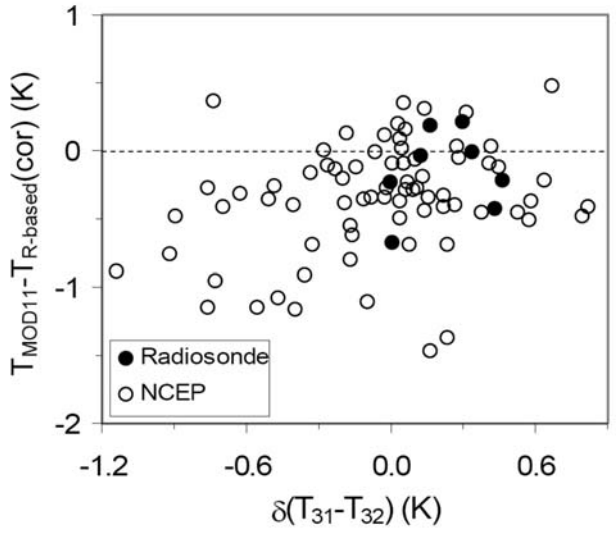

Figure 6. The same as in Figure 5 but $T_{R\text{-}based}$ corrected for $\delta(T_{31}-T_{32})$ effects due to mismatches in the atmospheric profiles using the linear regression from Figure 3.

conditions within the year 2004. The calculated $T_{R\text{-}based}$, the LST error ($\delta T = T_{M*D11} - T_{R\text{-}based}$) and the difference $\delta(T_{31}-T_{32})$ obtained from both radiosonde profiles and the NCEP profiles are shown in Table 6. If we compare the results obtained from radiosonde profiles and NCEP data, the LST errors and $\delta(T_{31}-T_{32})$ values are quite similar. Figure 7a shows the difference between R-based LSTs from radiosonde and NCEP profiles against the column water vapor difference $W_{rad}-W_{NCEP}$. Although some correlation ($R^2 = 0.48$) is observed, the LST differences cannot be fully explained by the water vapor differences. Figure 7b shows the LST differences against $\delta(T_{31}-T_{32})$, with a similar behavior as in Figure 4b and much larger correlation ($R^2 = 0.87$) than in Figure 7a. Thus, the $\delta(T_{31}-T_{32})$ differences can be a valuable means to asses the effect of mismatches in atmospheric profiles on the R-based LST calculations.

[50] According to Table 6, the LST error was within ±1.0 K for 53 of the total 64 Terra validation cases, and for 28 of the total 32 Aqua validation cases. These cases typically correspond to small absolute values of $\delta(T_{31}-T_{32})$. Large negative values of δT are usually associated with large negative values of $\delta(T_{31}-T_{32})$ (e.g., cases 14, 15, 17 and 34 with NCEP profiles). Figure 8a shows the LST errors against the $\delta(T_{31}-T_{32})$ values obtained from the radiosonde and NCEP profiles in Table 6. We can observe a similar relationship between δT and $\delta(T_{31}-T_{32})$ as for the Valencia test site (Figure 5), although the correlation is lower ($R^2 = 0.70$). It may be explained by the following reasons: (1) The LST and the atmospheric temperature profiles in the boundary layer vary in a much wider range at Hainich. (2) There are cloud pixels near the Hainich site in some cases and the effect of clouds on the retrieved LSTs is around 0.5 K.

[51] Since the surface emissivity of mixed forest at the Hainich site is expected to have similar uncertainty level as for the full-vegetation Valencia site, and the atmospheric water vapor does not exceed that of Valencia, we used the same $\delta(T_{31}-T_{32})$ thresholds to discriminate the atmospheric profiles that better represent the real atmospheric conditions. For 84 of the total 100 validation cases of Table 6, the $\delta(T_{31}-T_{32})$ values are within −0.3 and 0.5 K. Considering only these cases, the average LST bias is −0.30 K, the standard deviation is 0.50 K (RMSE = ±0.59 K), with LST errors between −1.7 and 0.7 K and 90% of the cases within ±1.0 K. The 16 outliers are marked with F in Table 6.

[52] As in section 4.2.2, the R-based LSTs can be corrected for the effect of mismatches in atmospheric profiles using the same correcting equation ($T_{R\text{-}based}(cor) = T_{R\text{-}based} + 2.4 \times \delta(T_{31}-T_{32}) - 0.2$). Figure 8b shows the difference between T_{MOD11} and $T_{R\text{-}based}(cor)$ as a function of $\delta(T_{31}-T_{32})$ for radiosonde and NCEP profiles. As in Figure 6, the LST errors show only small dependence on $\delta(T_{31}-T_{32})$. Thus, taking all the validation cases with the corrected R-based LSTs, the mean bias is −0.46 K, the standard deviation is 0.43 K, and RMSE = ±0.63 K, with LST errors ranging from −1.5 K to 0.4 K. These results are consistent with those for the cases passing the $\delta(T_{31}-T_{32})$ test but without applying the $\delta(T_{31}-T_{32})$ correction method.

[53] Since the Hainich test site is 60 km away from the Meiningen radiosounding station, there might be some error on the R-based LSTs due to possible atmospheric spatial variability. To evaluate this effect we selected a mixed forest site (50.54° N, 10.20° E) only 12 km from the Meiningen

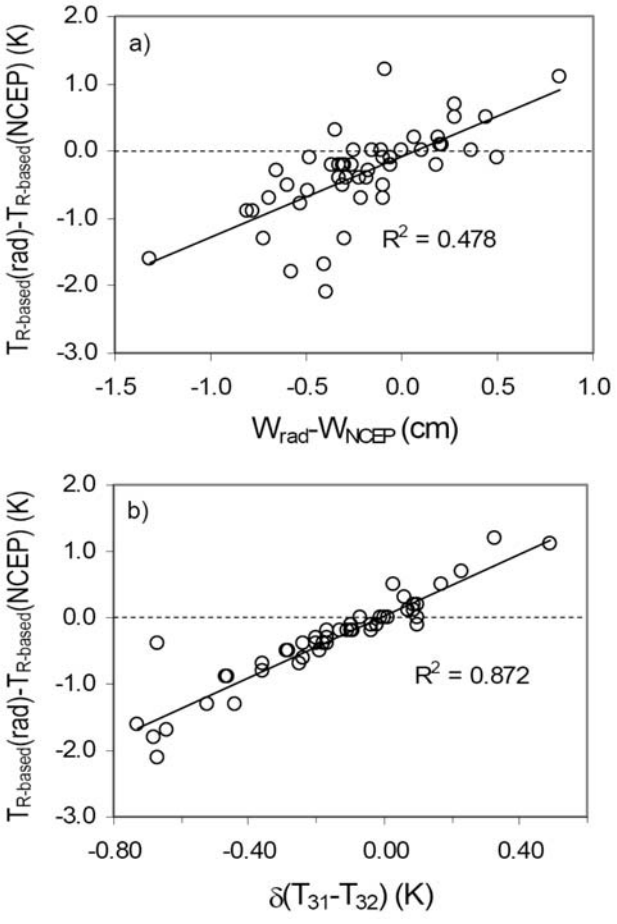

Figure 7. Difference between R-based LSTs obtained from radiosonde and NCEP profiles in the Hainich site (Terra and Aqua combined) against (a) column water vapor difference and (b) difference in $T_{31}-T_{32}$ simulated from local radiosonde and NCEP profiles.

Table 6. R-Based Validation Results for the Hainich Test Site With Radiosonde Profiles and NCEP Data for the Cases of Table 3[a]

	Radiosonde			NCEP		
Case	$T_{R\text{-based}}$ (°C)	δT (K)	$\delta(T_{31}-T_{32})$ (K)	$T_{R\text{-based}}$ (°C)	δT (K)	$\delta(T_{31}-T_{32})$ (K)
1	12.4	−0.2	0.37	12.6	−0.4	0.20
2	10.5	−0.9	0.25	10.9	−1.3	−0.42 F
3	19.6	−0.2	0.29	19.4	0.0	0.38
4	7.8	−0.2	−0.05	7.8	−0.2	−0.06
5	6.6	0.1	0.11	6.8	−0.1	0.01
6	15.9	−0.5	0.14	16.5	−1.1	−0.10
7	13.0	0.3	0.30	11.8	1.5	0.63 F
8	19.9	0.7	0.57 F	20.6	0.0	0.32
9	14.7	0.7	0.27	15.4	0.0	0.02
10	24.8	−1.7	−0.12	25.3	−2.2	−0.31 F
11	18.4	−0.7	0.41	17.9	−0.2	0.44
12	12.4	1.6	0.71 F	13.3	0.7	0.24
13	16.6	−0.6	0.05	15.5	0.5	0.54 F
14	13.8	−0.6	0.11	15.4	−2.2	−0.62 F
15	8.9	−1.5	−0.31 F	9.4	−2.0	−0.59 F
16	5.7	−0.1	0.10	5.7	−0.1	0.03
17	14.4	−0.5	0.10	16.2	−2.3	−0.58 F
18	6.5	0.1	0.24	6.4	0.2	0.31
19	6.9	−0.7	−0.06	7.2	−1.0	−0.23
20	1.9	−1.0	−0.08	2.0	−1.1	−0.18
21	11.1	0.6	0.47	11.3	0.4	0.38
22	0.7	−0.5	−0.02	0.9	−0.7	−0.12
23	4.4	−0.3	−0.03	4.6	−0.5	−0.16
24	11.1	0.5	0.36	11.0	0.6	0.45
25	12.1	0.0	0.05	12.5	−0.4	−0.12
26	14.6	−0.2	0.19	15.0	−0.6	0.01
27	21.9	−1.2	0.18	21.4	−0.7	0.35
28	12.9	−0.2	0.24	13.8	−1.1	−0.22
29	25.7	−0.6	0.16	25.5	−0.4	0.26
30	17.4	−0.7	−0.09	17.4	−0.7	0.01
31	15.4	−0.5	−0.01	15.4	−0.5	−0.02
32	21.6	0.1	0.36	22.3	−0.6	0.0
33	19.0	0.2	0.32	20.3	−1.1	−0.12
34	14.8	−0.3	−0.12	16.9	−2.4	−0.79 F
35	5.4	0.0	0.05	5.1	0.3	0.11
36	18.4	−0.7	0.10	18.8	−1.1	−0.14
37	3.1	−0.2	−0.06	3.2	−0.3	−0.10
38	20.3	0.2	0.32	20.3	0.2	0.33
39	10.0	−0.1	−0.05	10.1	−0.2	0.05
40	18.8	1.2	0.90 F	18.9	1.1	0.88 F
41	8.2	0.6	0.53 F	9.9	−1.1	−0.11
42	17.8	−0.5	0.17	18.2	−0.9	−0.03
43	7.9	−0.2	0.11	7.9	−0.1	0.11
44	14.8	1.0	0.94 F	16.1	−0.3	0.42
45	16.4	0.6	0.16	16.6	0.4	0.12
46	22.0	−0.9	0.25	21.3	−0.2	0.48
47	7.5	−0.4	0.29	8.0	−0.9	0.00
48	15.6	0.5	0.58 F	15.8	0.3	0.47
49	6.6	0.2	0.18	6.9	−0.1	−0.02
50	13.6	0.4	0.42	14.4	−0.4	0.06

[a] $\delta T = T_{M*D11} - T_{R\text{-based}}$. $\delta(T_{31}-T_{32})$ is the difference between the MODIS (actual) and the profile-based (calculated) brightness temperature difference in bands 31 and 32. F indicates cases failing the $\delta(T_{31}-T_{32})$ test. Cases 1–34 are from Terra. Cases 35–50 are from Aqua.

station, and used a new set of 30 clear-sky Terra scenes along 2004 for a comparison with the Hainich site. The same set of 30 radiosonde profiles were used in the R-based validation at these two sites. For the new site, the LST errors ($\delta T = T_{MOD11} - T_{R\text{-based}}$) yield mean bias of −0.31 K and standard deviation of 0.38 K, while for the Hainich site the mean bias is −0.27 K and the standard deviation is 0.52 K. The mean and standard deviation of $\delta(T_{31}-T_{32})$ was −0.02 and 0.28 K for the new site, and −0.01 and 0.31 K for the Hainich site. As the distance increases from 12 to 60 km, the standard deviations of δT and $\delta(T_{31}-T_{32})$ increase only slightly. This confirms the small horizontal spatial variation in the atmospheric temperature and water vapor profiles over a relatively flat area in clear sky situations, so we can use the Meiningen profiles at the Hainich site.

[54] Some of the new cases were classified as partial or full snow cover by the MODIS snow product (MOD02_L2). For these cases we used the spectral emissivity measured from a snow sample at the Mammoth Lake Cold Laboratory (http://www.icess.ucsb.edu/modis/EMIS/images/snowmam02.gif) (ε_{31} = 0.994 and ε_{32} = 0.986) instead of the vegetation emissivity values (ε_{31} = 0.979 and ε_{32} = 0.982). The effect of changing emissivity is shown in Figure 9, where the LST errors are plotted against $\delta(T_{31}-T_{32})$ for all cases (snow and snow-free) with appropriate emissivity values, and for the snow cases using the vegetation emissivity values instead. In Figure 9, snow-free cases follow approximately the typical relationship between δT and $\delta(T_{31}-T_{32})$. The snow cases using vegetation emissivity values clearly fall out of this relationship, increasing LST errors (in absolute value) by 0.8 K and $\delta(T_{31}-T_{32})$ values by 0.5 K when compared

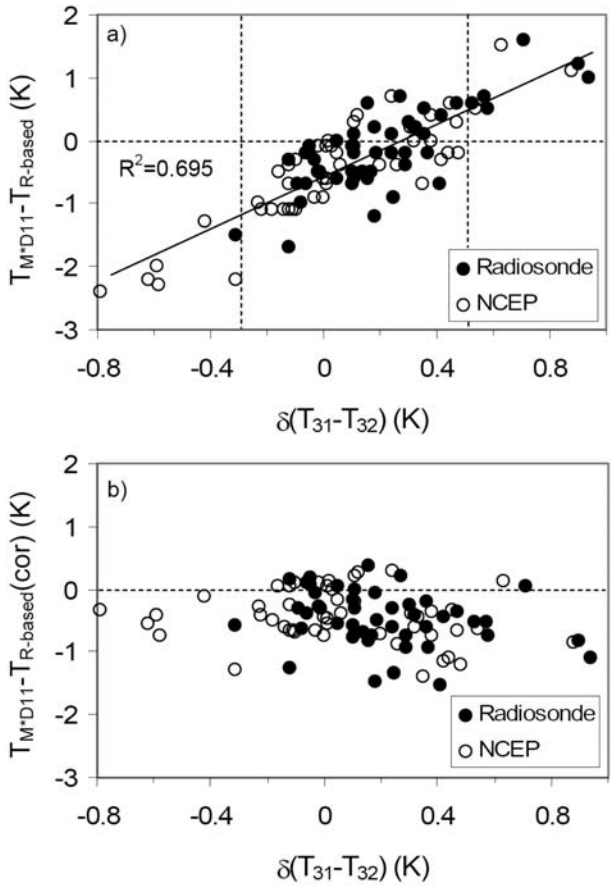

Figure 8. (a) M*D11_L2 LST minus R-based in situ LST calculated with radiosonde profiles and NCEP data against $\delta(T_{31}-T_{32})$ in the Hainich test site (Terra and Aqua combined). The solid line is the linear regression for all cases, and R^2 is the coefficient of determination. Vertical lines mark the valid $\delta(T_{31}-T_{32})$ range. (b) The same as Figure 8a but $T_{R\text{-based}}$ corrected for $\delta(T_{31}-T_{32})$ effects using the linear regression from Figure 3.

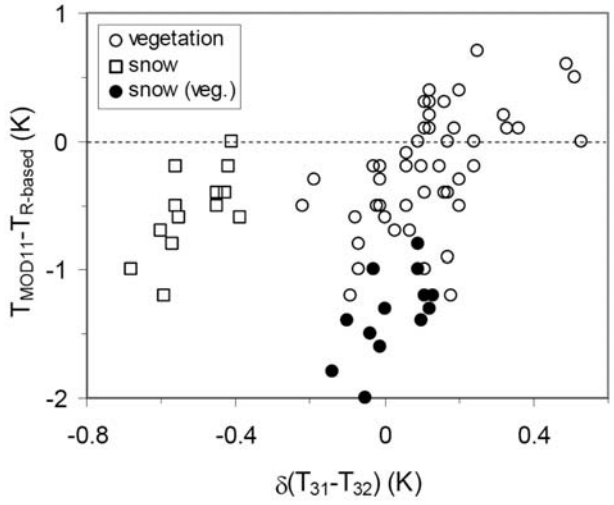

Figure 9. MOD11_L2 LST minus R-based in situ LST calculated with radiosonde profiles against $\delta(T_{31}-T_{32})$ for 30 new Terra scenes at Hainich and a forest site close to Meiningen. Open symbols represent vegetation and snow cases with appropriate emissivity values used for the R-based calculations. Closed circles represent snow cases with vegetation emissivity values.

with the same cases using snow emissivity values. For the snow cases, it appears that the assigned emissivity value in band 31 (0.994) is appropriate since it yields better LST errors, while the emissivity in band 32 (0.986) may be underestimated, thus resulting in large negative $\delta(T_{31}-T_{32})$ values. Recall that the snow emissivity used here correspond to flat snow sample and the emissivity of a forest covered by snow may be different due to surface structures. If ε_{32} was increased by 0.008 (then $\varepsilon_{31} = \varepsilon_{32}$), the $\delta(T_{31}-T_{32})$ values would be increased by approximately 0.5 K, thus bringing the snow case points in Figure 9 closer to the typical relationship between δT and $\delta(T_{31}-T_{32})$ for vegetation cases.

[55] These results give a good example for the high sensitivity of the R-based method to the variations in surface emissivities in bands 31 and 32. Therefore, we should be very careful in setting thresholds for the $\delta(T_{31}-T_{32})$ test and applying the $\delta(T_{31}-T_{32})$ correction method. It is necessary to work on a large data set of R-based validation results over a long period of days and to analyze the pattern in the δT versus $\delta(T_{31}-T_{32})$ distributions before setting the thresholds and determining the correction equations appropriately for different groups of cases. It is important to make sure that all the cases in each group form a good cluster, in which the average $\delta(T_{31}-T_{32})$ value corresponds to the effect of the errors in the assigned surface emissivities in bands 31 and 32, and the spread in $\delta(T_{31}-T_{32})$ around its average corresponds to the effect of mismatches in the atmospheric profiles.

[56] The R-based validation results of this section show the high accuracy of the M*D11_L2 LST product at the Hainich forest site (small biases and RMSE = ±0.6 K) in response to the great seasonal variations in LST, atmospheric temperature and column water vapor of the data set. However, the LST errors are smaller than those reported by *Wang et al.* [2008] for the same site. As pointed out in section 3.2, these authors applied the T-based method using single point, broadband longwave radiation measurements performed from a tower above the forest canopy. They used 95 nighttime data points from the beginning of 2004 to mid-2005, obtaining a mean bias of −2.21 K and RMSE = ±2.51 K. We believe that the largest part of these errors is due to the thermal heterogeneity of the forest area and the lack of representativity of the single-point measurements at the MODIS TIR pixel scale. A large error may also exist in the in situ LST estimation from broadband longwave radiation measurements made from high towers. In our opinion, the R-based method provides a more reliable validation approach than the T-based method for heterogeneous and complex surfaces such as the Hainich forest.

5. Conclusions

[57] The V5 level 2 MODIS LST product was validated over rice fields in Valencia, Spain and the Hainich forest in Germany. For the Valencia site, the study included conventional T-based validation using in situ LST measurements, and an alternative R-based validation using locally measured radiosonde profiles and NCEP tropospheric analysis data. Only the R-based validation was applied for the Hainich test site, with radiosonde profiles measured at a nearby radio sounding station and NCEP profiles. Since T-based validation relies on independently measured ground LSTs, it provides a direct evaluation of the MODIS LST retrieval algorithm. The results shown here for the Valencia test site (RMSE = ±0.5–0.6 K, with most of the cases in the ±1.0 K range) could be representative of the accuracy of the MOD11_L2 product over homogeneous areas in clear sky conditions. However, these results correspond to one single type of surface (nearly full cover of green vegetation, with stable, high emissivity), a limited temperature range (25–31°C), but quite variable atmospheric water vapor (1.7–3.8 cm).

[58] Due to the difficulties in scaling up from the ground, point LST measurements to the 1 km^2 MODIS pixel, only a few surface types are suitable for T-based validation within a useful uncertainty of ±1 K in the ground measured LSTs. Even in these circumstances, the collection of in situ LST measurements is a demanding task often limited to short-term, dedicated field campaigns. Therefore, the T-based method is not appropriate for the global validation of satellite derived LSTs. The R-based method [*Wan and Li*, 2008] provides an alternative since it does not require ground LST measurements. Validation areas should be homogeneous in emissivity (but not necessarily in temperature), a case which can be met quite frequently for the wavelength range of MODIS bands 31 and 32. The strongest limitation of the R-based method is the need for accurate atmospheric profiles and surface emissivity data to be used in radiative transfer calculations.

[59] The results of this paper demonstrate that the $\delta(T_{31}-T_{32})$ test, involving the difference between the actual MODIS and the profile-based calculated brightness temperature differences in bands 31 and 32, can be used to check the suitability of the atmospheric profiles used in the R-based validation. For the Valencia test site, the R-based method was applied with local radiosonde measurements for eight cases, and with NCEP profiles for 74 cases

including daytime and nighttime data. We obtained that the cases with the small LST errors were associated with $-0.3\ \mathrm{K} < \delta(T_{31}-T_{32}) < 0.5\ \mathrm{K}$. All the eight cases with local radiosonde data and 49 cases with NCEP profiles met this condition, for which the RMSE was ±0.5–0.6 K and LST errors were within ±1.0 K for most of the cases. These error figures are similar to the results obtained in the T-based validation. They are also in agreement with those shown by *Wan* [2008] for different surface types, temperatures and atmospheric water vapor contents.

[60] A similar analysis was applied to the Hainich validation data set, which covers larger ranges of LST (0–25°C) and atmospheric water vapor (0.5–2.6 cm). For the cases meeting the above $\delta(T_{31}-T_{32})$ condition, the LST errors were comparable (RMSE = ±0.6 K). Thus, the $\delta(T_{31}-T_{32})$ test appears as a valuable means to select accurate atmospheric profiles for the R-based validation, but may be also sensitive to the presence of thin, undetected cirrus clouds and aerosols above the average loading.

[61] In essence, the R-based validation approach is the radiative transfer method to retrieve LST from single band data plus the $\delta(T_{31}-T_{32})$ test or $\delta(T_{31}-T_{32})$ correction. The results of this paper confirmed that the R-based method can be used to validate the MODIS LST product at an accuracy better than 1 K (mean bias in range −0.3 K to −0.5 K and standard deviation around 0.5 K statistically) at fully vegetated sites. This is close to the accuracy (better than 1 K and within 0.5 K in most cases) capably achieved by the MODIS LST product in situations with well known surface emissivities such as lakes and fully vegetated regions under dry to wet atmospheric conditions, as demonstrated in the simulations during the development of the algorithm and indicated in the previous validation studies.

[62] The R-based method using appropriate, nearly concurrent measured radiosonde profiles and global NCEP atmospheric data opens the possibility for the semioperational validation and diagnostics of the MODIS LST product at global scale on a routine basis. It could be applied over surfaces where the ground LST measurement is not feasible (forests, surfaces with partial vegetation cover, semiarid areas, deserts, remote regions, etc.), and along extended periods with varied temperature regimes and atmospheric conditions. Accurate measurements of surface emissivities at some sites may be necessary. A limitation may exist in the case of wet atmospheres where the atmospheric effect is stronger and subject to larger uncertainties. Homogeneous and stable atmospheric conditions are required, so the profile used in the R-based validation can represent adequately the atmospheric state at the time of the MODIS observation. The $\delta(T_{31}-T_{32})$ test is recommended to check the suitability of atmospheric profiles for each validation case in long periods of time at each site. The $\delta(T_{31}-T_{32})$ correction method has the ability to correct the $\delta(T_{31}-T_{32})$ dependence in estimated LST errors for the effect of mismatches in atmospheric profiles so that the accuracy and suitability requirements for the atmospheric profile may be less strict. This alternative method should be evaluated at more validation sites in the future.

[63] **Acknowledgments.** This work was funded by the Spanish Ministerio de Educación y Ciencia (CGL2007-29819-E and CGL2007-64666/CLI and research grant of J. M. Galve) and the Conselleria d'Educació of the Generalitat Valenciana (PROMETEO/2009/086). Z. Wan was supported by contract NNG04HZ15C and cooperative agreement NNX08AE62A of NASA. We thank the University of Valencia Thermal Remote Sensing Unit for the ground LST measurements and the Centro de Estudios Ambientales del Mediterraneo (CEAM) for the radiosonde data at the Valencia site.

References

Barnes, W. L., T. S. Pagano, and V. V. Salomonson (1998), Prelaunch characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1, *IEEE Trans. Geosci. Remote Sens.*, *36*, 1088–1100, doi:10.1109/36.700993.

Berk, A., G. P. Anderson, P. K. Acharya, J. H. Chetwynd, L. S. Bernstein, E. P. Shettle, M. W. Matthew, and S. M. Adler-Golden (1999), *MODTRAN 4 User's Manual*, 95 pp., Air Force Res. Lab., Space Vehicles Dir., Air Force Mater. Command, Hanscom Air Force Base, Mass.

Bugbee, B., M. Droter, O. Monje, and B. Tanner (1998), Evaluation and modification of commercial infra-red transducers for leaf temperature measurement, *Adv. Space Res.*, *22*(10), 1425–1434, doi:10.1016/S0273-1177(98)00213-0.

Coll, C., V. Caselles, J. M. Galve, E. Valor, R. Niclòs, J. M. Sánchez, and R. Rivas (2005), Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data, *Remote Sens. Environ.*, *97*, 288–300, doi:10.1016/j.rse.2005.05.007.

Coll, C., V. Caselles, J. M. Galve, E. Valor, R. Niclòs, and J. M. Sánchez (2006), Evaluation of split-window and dual-angle correction methods for land surface temperature retrieval from Envisat/AATSR data, *J. Geophys. Res.*, *111*, D12105, doi:10.1029/2005JD006830.

Coll, C., V. Caselles, E. Valor, R. Niclòs, J. M. Sánchez, J. M. Galve, and M. Mira (2007), Temperature and emissivity separation from ASTER data for low spectral contrast surfaces, *Remote Sens. Environ.*, *110*, 162–175, doi:10.1016/j.rse.2007.02.008.

Coll, C., S. J. Hook, and J. M. Galve (2009), Land surface temperature from the Advanced Along-Track Scanning Radiometer: Validation over inland waters and vegetated surfaces, *IEEE Trans. Geosci. Remote Sens.*, *47*(1), 350–360, doi:10.1109/TGRS.2008.2002912.

Dash, P., and A. Ignatov (2008), Validation of clear-sky radiances over oceans simulated with MODTRAN4.2 and global NCEP GDAS fields against nighttime NOAA15–18 and MetOp-A AVHRR data, *Remote Sens. Environ.*, *112*(6), 3012–3029, doi:10.1016/j.rse.2008.02.013.

French, A. N., T. J. Schmugge, J. C. Ritchie, A. Hsu, F. Jacob, and K. Ogawa (2008), Detecting land cover change at the Jornada Experimental Range, New Mexico with ASTER emissivities, *Remote Sens. Environ.*, *112*, 1730–1748, doi:10.1016/j.rse.2007.08.020.

Hook, S. J., R. G. Vaughan, H. Tonooka, and S. G. Schladow (2007), Absolute radiometric in-flight validation of mid infrared and thermal infrared data from ASTER and MODIS on the Terra spacecraft using the Lake Tahoe, CA/NV, USA, automated validation site, *IEEE Trans. Geosci. Remote Sens.*, *45*, 1798–1807, doi:10.1109/TGRS.2007.894564.

Kalnay, E., et al. (1996), The NCEP/NCAR 40 year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Merchant, C. J., and P. Le Borgne (2004), Retrieval of sea surface temperature from space, based on modeling of infrared radiative transfer: Capabilities and limitations, *J. Atmos. Oceanic Technol.*, *21*, 1734–1746, doi:10.1175/JTECH1667.1.

Momeni, M., and M. Saradjian (2007), Evaluating NDVI-based emissivities of MODIS bands 31 and 32 using emissivities derived by day/night LST algorithm, *Remote Sens. Environ.*, *106*(2), 190–198, doi:10.1016/j.rse.2006.08.005.

Ribeiro da Luz, B., and J. K. Crowley (2007), Spectral reflectance and emissivity features of broad leaf plants: Prospects for remote sensing in the thermal infrared (8.0–14.0 μm), *Remote Sens. Environ.*, *109*, 393–405, doi:10.1016/j.rse.2007.01.008.

Rubio, E., V. Caselles, C. Coll, E. Valor, and F. Sospedra (2003), Thermal-infrared emissivities of natural surfaces: Improvements on the experimental set-up and new measurements, *Int. J. Remote Sens.*, *24*(24), 5379–5390, doi:10.1080/0143116031000102412.

Salomonson, V. V., W. L. Barnes, P. W. Maymon, H. Montgomery, and H. Ostrow (1989), MODIS: Advanced facility instrument for studies of the Earth as a system, *IEEE Trans. Geosci. Remote Sens.*, *27*, 145–153, doi:10.1109/36.20292.

Seemann, S. W., E. E. Borbas, J. Li, W. P. Menzel, and L. E. Gumley (2006), MODIS atmospheric profile retrieval Algorithm Theoretical Basis Document, 40 pp., Univ. of Wis.–Madison, Madison. (Available at http://modis-atmos.gsfc.nasa.gov/_docs)

Sicard, M., P. R. Spyak, G. Brogniez, M. Legrand, N. K. Abuhassan, C. Pietras, and J. P. Buis (1999), Thermal infrared field radiometer for vicarious cross-calibration: Characterization and comparisons with other field instruments, *Opt. Eng.*, *38*(2), 345–356, doi:10.1117/1.602094.

Snyder, W., Z. Wan, Y. Zhang, and Y.-Z. Feng (1998), Classification-based emissivity for land surface temperature measurement from space, *Int. J. Remote Sens.*, *19*, 2753–2774, doi:10.1080/014311698214497.

Trigo, I. F., I. T. Monteiro, F. Olesen, and E. Kabsch (2008), An assessment of remotely sensed land surface temperature, *J. Geophys. Res.*, *113*, D17108, doi:10.1029/2008JD010035.

Wan, Z. (2008), New refinements and validation of the MODIS land-surface temperature/emissivity products, *Remote Sens. Environ.*, *112*, 59–74, doi:10.1016/j.rse.2006.06.026.

Wan, Z., and J. Dozier (1996), A generalized split-window algorithm for retrieving land surface temperature from space, *IEEE Trans. Geosci. Remote Sens.*, *34*(4), 892–905, doi:10.1109/36.508406.

Wan, Z., and Z.-L. Li (1997), A physics-based algorithm for retrieving land surface emissivity and temperature from EOS/MODIS data, *IEEE Trans. Geosci. Remote Sens.*, *35*, 980–996, doi:10.1109/36.602541.

Wan, Z., and Z.-L. Li (2008), Radiance-based validation of the V5 MODIS land-surface temperature product, *Int. J. Remote Sens.*, *29*(17–18), 5373–5395, doi:10.1080/01431160802036565.

Wan, Z., Y. Zhang, Q. Zhang, and Z.-L. Li (2002), Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data, *Remote Sens. Environ.*, *83*, 163–180, doi:10.1016/S0034-4257(02)00093-7.

Wan, Z., Y. Zhang, Q. Zhang, and Z.-L. Li (2004), Quality assessment and validation of the MODIS global land surface temperature, *Int. J. Remote Sens.*, *25*(1), 261–274, doi:10.1080/0143116031000116417.

Wang, W., S. Liang, and T. Meyers (2008), Validating MODIS land surface temperature products using long-term nighttime ground measurements, *Remote Sens. Environ.*, *112*, 623–635, doi:10.1016/j.rse.2007.05.024.

C. Coll and J. M. Galve, Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, Dr. Moliner, 50, E-46100 Burjassot, Spain. (cesar.coll@uv.es)

Z. Wan, ICESS, University of California, Santa Barbara, CA 93106, USA. (wan@icess.ucsb.edu)